Advanced Configuration for the **Dell PowerConnect 5316M Blade Server Chassis Switch**

BY SURENDRA BHAT

SAURABH MALLIK

Enterprises can take advantage of advanced configuration options for the Dell™ PowerConnect™ 5316M Gigabit Ethernet switch for the Dell PowerEdge™ blade server chassis to help optimize network access for blade servers.

Traditional network management methods have typically emphasized bandwidth and reliability. As network complexity increases, however, administrators also require advanced management options to help ensure quality and efficiency. The 16-port Dell PowerConnect 5316M Gigabit Ethernet switch for the Dell PowerEdge blade server chassis includes features such as load balancing using Multiple Spanning Tree Protocol (MSTP), quality-of-service (QoS) management for traffic priority, and Internet Group Management Protocol (IGMP) snooping. This article provides a general overview of these features and examples of their implementation.

Related Categories:

Dell PowerConnect switches

Dell PowerEdge blade servers

Ethernet

Network fabric

Visit www.dell.com/powersolutions for the complete category index.

Load balancing with Multiple Spanning Tree Protocol

Understanding MSTP first requires understanding the basic Spanning Tree Protocol (STP) for link management, defined in the IEEE 802.1d standard. A Layer 2 switch builds a forwarding table by learning the Media Access Control (MAC) addresses of the frames entering the switch. If it does not find the destination MAC address for a particular frame in the MAC address table, it forwards the frame on all ports other than the source port—meaning that if any loops exist in the Layer 2 network, frames may be forwarded endlessly and cause broadcast storms.

STP is designed to provide path redundancy while preventing these types of network loops. It creates a tree topology among the various network components by designating a root bridge for the entire network and another bridge for each Ethernet segment, which it accomplishes by exchanging Bridge Protocol Data Units (BPDUs) between the switches after considering attributes such as bridge and port priority, MAC address, and link bandwidth. To help avoid loops, some ports are placed into the blocking state for all data while they continue receiving BPDUs. If the network topology changes, the blocking ports move through the intermediate listening and learning states before settling in the forwarding state.

Although traditional STP does help prevent Layer 2 forwarding loops in a general network topology, convergence can take up to 30–60 seconds. Rapid Spanning Tree Protocol (RSTP), defined in the IEEE 802.1w standard, alters the port roles and takes advantage of point-to-point wiring to provide rapid convergence of the spanning tree.

STP and RSTP maintain a single topology instance for all virtual LANs (VLANs). VLANs create multiple logical networks that can be grouped by function or application rather than physical location, and create separate broadcast domains equivalent to IP subnetworks. Using STP and RSTP with VLANs can result in suboptimal paths for certain traffic. However, MSTP, defined in the IEEE 802.1s standard, can maintain multiple spanning tree instances and assign VLANs to those instances, allowing the use of different physical paths for different VLAN traffic and helping create efficient load balancing for network resources.

When MSTP is enabled on the PowerConnect 5316M switch, the switch uses RSTP to develop a loop-free topology for each spanning tree instance. Administrators can enable MSTP on the switch with the following command:

```
console(config)# spanning-tree mode mstp
```

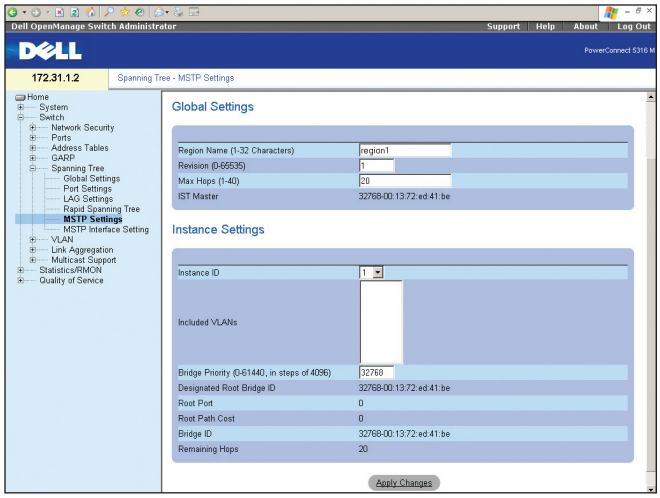

Figure 1. *MSTP Settings screen for the Dell PowerConnect 5316M switch*

Figure 1 shows the MSTP Settings screen in the Dell OpenManage™ Switch Administrator Web browser–based graphical user interface (GUI) for the PowerConnect 5316M. Administrators can use this screen to configure the Multiple Spanning Tree (MST) region name, revision number, and VLAN-to-instance mapping. Adjacent switches must use the same settings for them to become part of the same MST region. Administrators can change the bridge priority setting to influence the root bridge selection for a particular spanning tree instance.

Administrators can use the following commands to manipulate the cost and priority settings, which play a major role in determining the active and blocked interfaces in the topology:

```
console(config-if)# spanning-tree mst
   instance_id cost cost
console(config-if)# spanning-tree mst
   instance_id port-priority priority
```

They can execute these commands from the Ethernet or link aggregation group (LAG) interface configuration mode, and can do so separately for each MSTP instance. A LAG interface represents the link aggregation of multiple physical interfaces between two Ethernet switches and appears to the spanning tree algorithm as a single logical interface.

Using quality of service for traffic management

In practical situations, traffic from blade servers to a chassis switch may consist of more than the uplink (traffic from the chassis switch to the external network): the switch accommodates surplus traffic in queues that act as buffers, which results in network latency and can cause problems for time-sensitive data such as voice over IP (VoIP) traffic. *Quality of service* refers to traffic mechanisms that prioritize data flow for certain data sources over others, in contrast to a "best-effort" mechanism that allocates network resources equally to competing applications. QoS in a Layer 2 switch, such as the PowerConnect 5316M, can thus be referred to as a priority queuing scheme.

Figure 2 depicts the three phases a packet passes through when it enters the switch: classification, queuing, and scheduling. *Classification* is the process of distinguishing different types of traffic by examining the packet fields. The PowerConnect 5316M can detect a packet's priority through two schemes: class of service (CoS) and differentiated services code point (DSCP). CoS is part of the IEEE 802.1p standard and is contained in a three-bit field in the VLAN header; its values range from 0 to 7, with 7 having the highest priority. DSCP is an eight-bit field in the Layer 3 IP header; its values range from 0 to 63, with 63 having the highest priority. Default CoS values can be assigned to the incoming frames on each port. Administrators can set the trust mode on the PowerConnect 5316M, which carries forward the embedded priority in the incoming frame, to use either CoS or DSCP with the following command (where *scheme* is either dscp or cos):

```
console(config)# qos trust scheme
```

Queuing is the process of assigning incoming packets to the appropriate egress queue. The switch bases the outgoing port on the forwarding table entry for the destination MAC address. Each port has four queues, and the packet is assigned to one of these based on the priority tag. The following commands map CoS and DSCP values to specific queues:

```
console(config)# wrr-queue cos-map queue_id
   cos_values_list
console(config)# qos map dscp-queue
   dscp_values_list to queue_id
```

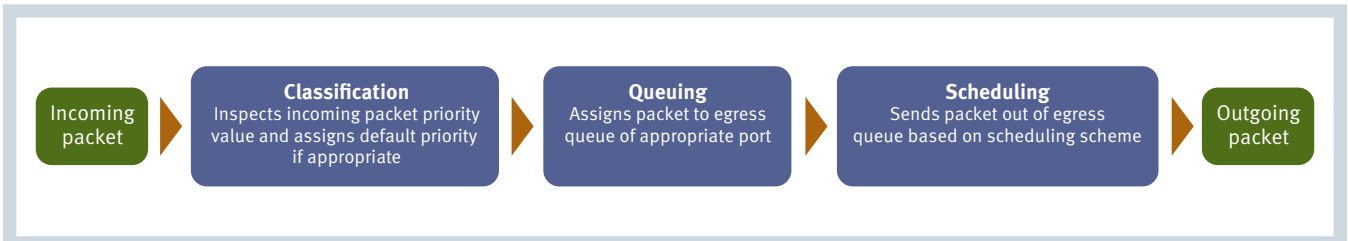

Figure 2. *QoS flow diagram for the Dell PowerConnect 5316M switch*

"As network complexity increases, administrators require advanced management options to help ensure quality and efficiency."

Scheduling services the queues based on their configured priorities. Queues can be configured in one of two ways:

- **Strict priority (SP):** SP helps ensure that the highest-priority queues are serviced first and that critical, time-sensitive traffic is prioritized over less critical or less time-sensitive traffic.
- **Weighted round robin (WRR):** WRR helps ensure that a single high-priority application does not take over all bandwidth by allocating queues proportionally to their assigned weights.

Figure 3 shows the QoS Global Queue Settings screen in the Dell OpenManage Switch Administrator GUI for the PowerConnect 5316M. The SP and WRR settings apply to all ports for egress traffic.

Snooping with Internet Group Management Protocol

IGMP is the standard protocol for IP multicasting in a broadcast domain. Multicast groups are segregated host groups that share packets among themselves; multicast datagrams are transmitted only to multicast group members. The router maintains group membership lists through IGMP: for example, a host joining a group sends an IGMP join message to the router, or, alternatively, the switch forwards IGMP queries from the router

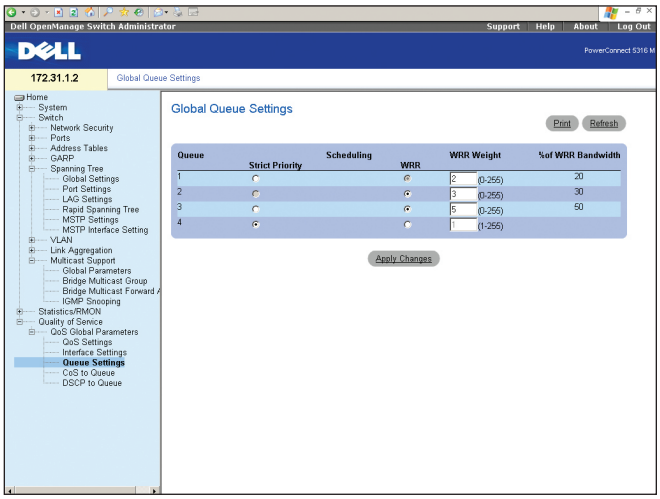

Figure 3. *QoS Global Queue Settings screen for the Dell PowerConnect 5316M switch*

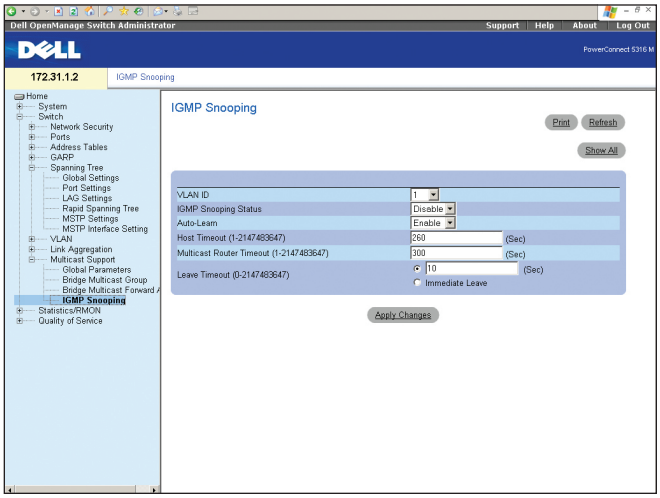

Figure 4. *IGMP Snooping screen for the Dell PowerConnect 5316M switch*

to all the ports connected to it. The advantage of the latter scheme is that it allows a single copy to be sent to the switch, which then creates copies and transmits them. By default, a Layer 2 switch handles IP multicast traffic in the same manner as broadcast traffic, by forwarding it to all interfaces. Thus, IGMP traffic can potentially flood the interfaces and cause network congestion.

Layer 2 switches can use a feature called IGMP snooping to help prevent this flooding. With this feature enabled, the switching application-specific integrated circuit (ASIC) forwards all IGMP packets to the switch processor, which then analyzes the incoming packets, maps the ports to multicast groups, and determines which ports will send out the IGMP queries and which routing protocols to use. This method uses a multicast filtering database; the switch checks this database and only forwards multicast datagrams to member interfaces, helping optimize interface and bandwidth usage.

To use this feature on the PowerConnect 5316M, administrators must enable both the bridge multicast filtering feature and IGMP snooping. They can do this in the Dell OpenManage Switch Administrator GUI from the Multicast Global Parameters screen and the IGMP Snooping screen (see Figure 4), or by using the following commands:

```
console(config)# bridge multicast filtering
console(config)# ip igmp snooping
```

Removing hosts from groups requires updating the filtering database, which can be done immediately upon receiving an IGMP leave command or after a set leave time-out period. Administrators can configure this setting from the IGMP Snooping screen or with the following command (where *setting* is either `time-out` or `immediate-leave`):

```
console(config-if)# ip igmp snooping
   leave-time-out setting
```

Implementing Dell PowerConnect 5316M switches

Two examples can help administrators understand how to implement advanced configuration options for PowerConnect 5316M switches: one for MSTP and one for QoS.

Example MSTP implementation

Administrators can use MSTP with PowerConnect 5316M switches to enable load balancing for a blade server chassis. Figure 5 shows a typical high-availability implementation using two PowerConnect 5316M chassis switches along with two external PowerConnect 5324 switches that provide redundancy and availability in case of a switch failure. MSTP is used to balance the blade traffic from the PowerConnect 5316M switches to the two PowerConnect 5324 switches.

For example purposes, assume that the network includes VLANs 2–7, with VLANs 2–4 assigned to MSTP instance 1 and VLANs 5–7 assigned to MSTP instance 2. Link aggregation between the PowerConnect 5316M and PowerConnect 5324 switches groups together multiple links and helps provide load balancing and fault tolerance. (These configuration examples still apply when using individual physical interfaces rather than LAGs.)

To simplify the example, consider the connectivity of only the PowerConnect 5316M switch shown in the lower right corner of Figure 5. The cost and priority settings for this switch's LAG 1 and LAG 2 interfaces can be adjusted separately for MSTP instances 1 and 2 to set the LAG port roles and states shown in Figure 6. These port roles and states are defined in the RSTP standard; the root port provides the lowest cost when the switch forwards packets to the root switch and would be in the forwarding state, while the alternate port offers a different path to the root switch and would be in the discarding state. Figures 7 and 8 show the resulting topologies for the two MSTP instances.

If this network had used traditional STP, only one of the LAGs would carry traffic for all the VLANs. Because MSTP allows multiple forwarding paths for data traffic, the network can maintain separate logical topologies for different VLAN groups to help make efficient use of network resources. A similar setup can also be used with the other PowerConnect 5316M switch.

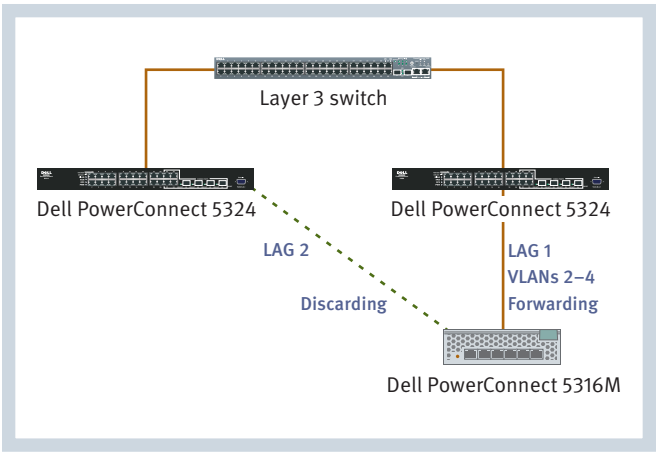

Figure 7. *Network topology for MSTP instance 1 in the example high-availability MSTP implementation*

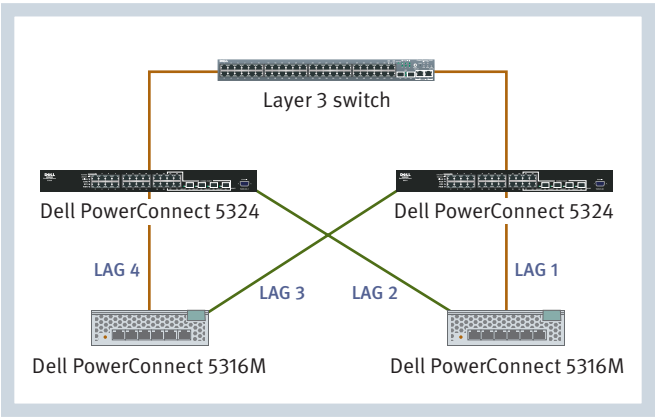

Figure 5. *Example high-availability MSTP implementation using Dell PowerConnect 5316M and PowerConnect 5324 switches*

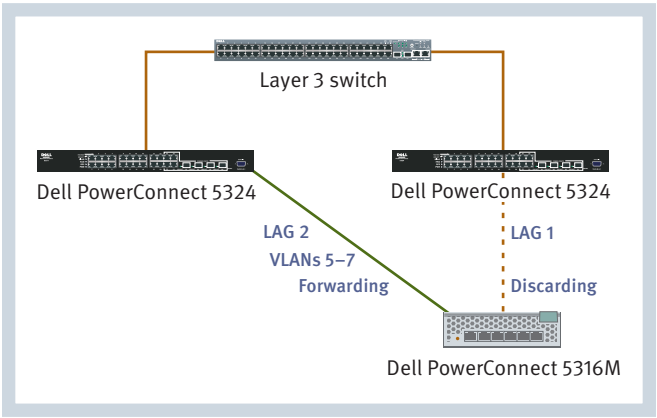

Figure 8. *Network topology for MSTP instance 2 in the example high-availability MSTP implementation*

MSTP instance	VLANs	LAG 1 role	LAG 1 state	LAG 2 role	LAG 2 state
1	2–4	Root	Forwarding	Alternate	Discarding
2	5–7	Alternate	Discarding	Root	Forwarding

Figure 6. *LAG port roles and states for a Dell PowerConnect 5316M switch in the example MSTP implementation*

Port number	CoS value
1	7
2	7
3	5
4	5
5	3
6	3
7	3
8	1
9	1
10	1

Figure 9. *Example default CoS values used to assign incoming traffic priorities for Dell PowerConnect 5316M internal ports*

Example QoS implementation

Implementing the QoS feature in blade server chassis switches may be necessary to prioritize the traffic from certain blades. For example, to assign priorities to incoming traffic coming from each of 10 blades, administrators could assign default CoS values to the 10 internal ports of the chassis switch that connects to the blade LAN on Motherboards (LOMs), as shown in Figure 9.

Administrators could also assign default priority values for incoming traffic on the six external ports, but typically they should designate all of these ports with a common priority that allocates traffic to the best-effort queue. They should enable the trust mode on all 16 ports so that if incoming packets arrive with a CoS or DSCP value, the switch does not overwrite these values with the default port setting.

CoS value	Queue
0	1
1	1
2	2
3	2
4	3
5	3
6	4
7	4

Figure 10. *Example CoS values mapped to queue numbers for the Dell PowerConnect 5316M switch*

Queue	Priority configuration	WRR weight	WRR bandwidth
1	WRR	2	20%
2	WRR	3	30%
3	WRR	5	50%
4	SP	N/A	N/A

Figure 11. *Example configuration using the Global Queue Settings screen for the Dell PowerConnect 5316M switch*

For each egress port, the SP queues are defined starting from the higher queue numbers; therefore, the higher CoS values are mapped to higher queue numbers, as shown in Figure 10. The queues can be configured using the parameters on the Global Queue Settings screen of the Dell OpenManage Switch Administrator GUI, as shown in Figure 11. Queue 4 is defined as an SP queue, and weights are assigned to the three WRR queues. With these settings, the traffic from blades 1 and 2 is assigned to the SP queue, while traffic from the other blades is assigned using the WRR algorithm to help ensure higher priority for higher CoS values.

Optimizing network traffic for blade servers

The advanced configuration options for the Dell PowerConnect 5316M blade server chassis switch include MSTP for load balancing, QoS for priority management, and IGMP snooping for IP multicast traffic. Using these capabilities enables administrators to optimize blade server data traffic and enhance deployment flexibility in network infrastructures.

Surendra Bhat *is a test engineer senior analyst in the Enterprise System Test Group at the Dell Bangalore Development Center. He has a bachelor's degree in Electronics Engineering from the University of Mumbai. His interests include networking and storage technologies.*

Saurabh Mallik *is a test engineer analyst in the Enterprise System Test Group at the Dell Bangalore Development Center. He has a bachelor's degree in Electronics and Communications Engineering from the Peoples Education Society Institute of Technology (PESIT).*

more online

www.dell.com/**powersolutions**

QUICK LINK

Dell PowerConnect 5316M User's Guide:
support.dell.com/support/edocs/network/PC5316M